DOCUMENT RESUME

ED 351 356 TM 019 165

AUTHOR Byrne, Barbara M.; And Others
TITLE Gender Differences in Adolescent Depression: Testing for Invariant Measurement and Structure for the BDI (French Version).
PUB DATE Jul 92
NOTE 31p.; Paper presented at the International Congress of Psychology (25th, Brussels, Belgium, July 1992).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Adolescents; *Depression (Psychology); *Diagnostic Tests; *Factor Structure; Females; Foreign Countries; French; Males; Psychological Testing; Psychometrics; *Sex Differences; Test Format; *Test Validity
IDENTIFIERS *Beck Depression Inventory; Canada; French (Canadian)

ABSTRACT

A study of the Beck Depression Inventory (BDI) was conducted to: (1) test for the factorial validity of the French version of the BDI (BDI-FR) separately for 551 non-clinical Francophone adolescent males and 601 non-clinical Francophone adolescent females; (2) cross-validate findings across a second independent sample for each sex; and (3) test for equivalent factorial structure across gender for this population. Except for 1 differential factor loading for females (item 17), the hypothesized higher-order 3-factor structure was found to be tenable across validation samples and factorially invariant across gender. Although the presence of error covariances for females partially replicated a similar pattern for English female adolescents, results generally provide strong support for the psychometric soundness of the BDI-FR. Two tables and three figures are included. (Author/SLD)

Gender Differences in Adolescent Depression: Testing for Invariant Measurement and Structure for the BDI (French Version)

Barbara M. Byrne

Pierre Baron T. Leanne Campbell

University of Ottawa

Paper presented at the XXV International Congress of Psychology, Brussels, July 1992

Abstract

The purposes of the study were threefold: (a) to test for the factorial validity of the French version of the BDI (BDI-FR; Bourque & Beaudette, 1982) separately for nonclinical adolescent Francophone males (n=551) and females (n=601), (b) to cross-validate findings across a second independent sample for each sex, and (c) to test for equivalent factorial structure across gender for this population. Except for one differential factor loading for females (Item 17), the hypothesized higher-order 3-factor structure was found to be both tenable across validation samples, and factorially invariant across gender. Although the presence of error covariances for females partially replicated a similar pattern for English female adolescents, results generally provide strong support for the psychometric soundness of the BDI-FR.

Gender Differences in Adolescent Depression: Testing for Invariant Measurement and Structure for the BDI (French Version)

The present study extends the work of Byrne and colleagues (Byrne & Baron, in press a, in press b; Byrne, Campbell, & Baron, 1991) in validating . . hierarchical factorial structure of the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) for nonclinical adolescents. In particular, it focuses on the French version of the instrument (BDI-FR) and tests for (a) factorial validity based on separate samples of French speaking male and female adolescents, (b) the replication of factorial structure across a second independent sample for each sex, and (c) the equivalency of factorial structure across gender for this population.

Based on confirmatory factor analyses (CFAs) involving three independent samples of nonclinical adolescents, Byrne and Baron (in press a) concluded that an hierarchical 3-factor structure of the BDI best represented the data. Specifically, their cross-validated findings demonstrated strong support for a second-order structure consisting of one higher-order general factor of depression, and three lower-order factors which they labelled Negative Attitude, Performance Difficulty, and Somatic Elements. Although this study represented the most stringent to date in

testing the validity of the BDI for use with normal adolescents, it followed the trend of previous factor analytic research by targeting English speaking samples. (for an extensive review of this literature, see Byrne & Baron, in press a).

Less well known, however, is a French version of the BDI (BDI-FR) developed and validated by Bourque and Beaudette (1982) for use with French speaking adults. Until the recent study by Byrne and Baron (in press b), the work of Baron and LaPlante (1984) represented the only known attempt to validate this instrument with French speaking adolescents. Addressing limitations in the Baron and LaPlante study, Byrne and Baron conducted a CFA study of the BDI-FR that included cross-validation based on three independent samples.

Although findings from the Byrne and Baron (in press b) study concurred with those based on the English version of the BDI (Byrne & Baron, in press a) by concluding an hierarchical 3-factor structure to most accurately represent the data, a differential pattern of factor loadings related to eight items was found for French adolescents; these were as follows: (a) Items 3 (failure), 5 (guilt), 7 (self-dislike), and 14 (self-image) loaded on Performance Difficulty, rather than on the Negative Attitude factor, (b) Items 4 (dissatisfaction) and 12 (withdrawal) loaded on Negative Attitude, rather than on the

Performance Difficulty factor, and (c) Items 17 (fatigue) and 20 (hypochondria) loaded on the Somatic Elements, rather than on the Performance Difficulty factor. A schematic representation of this factorial structure is shown in Figure 1.

Insert Figure 1 about here

While the factorial validity of a measuring instrument is crucial to the credibility of findings from substantive research involving its use, it is equally critical that such validity be group-invariant in the conduct of multigroup comparison research. An important assumption underlying such endeavours is that the instrument of measurement is factorially equivalent across groups. More specifically, this type of research assumes the equivalency of the assessment measure with respect to (a) number of underlying factors, (b) item measurements (i.e., all item content is equivalently interpreted by both groups), and (c) theoretical structure (i.e, pattern of relations among the underlying factors is the same across groups). These questions bear on the construct validity of the intrument. Despite the serious importance of factorial invariance, however, such information is most often assumed in substantive multigroup research and rarely, if ever, directly tested.

Addressing this issue, Byrne et al. (1991) tested these equality assumptions across adolescent males and females as they related to the English version of the BDI. Although they reported the BDI, for the most part, to be factorially invariant across gender, their findings revealed Item 20 (hypochondria) to load differentially for males and females, and Items 8 (self-accusation) and 10 (crying) to be noninvariant. Whereas Item 20 loaded onto the Performance Difficulty factor for males, it loaded on Negative Attitude for females. One additional gender difference identified in this research was the strong error covariance between Items 20 and 21 (libido) for females; no error covariances were evidenced for males.

The above study suggests that although the underlying factorial structure of the BDI appears to be most adequately represented by a higher-order structure that includes three first-order factors (Negative Attitude, Performance Difficulty, Somatic Elements) and one second-order factor of General Depression, for both English and French speaking nonclinical adolescents, there is some evidence of slight gender differences in the interpretation of BDI items. The present study, then, extends this previous construct validity work by testing assumptions bearing on the equivalency of factorial structure for the BDI-FR across French-speaking adolescent males and females.

Method

Sample and Procedures

The data comprised BDI-FR responses from 1152 (551 males; 601 females) French Canadian adolescents. All subjects were high school students (grades 9-12) from two large urban areas in central Canada. Only questionnaires with complete data were included in the analyses.

Subjects completed the BDI, along with other assessment measures, during one regular class period; all testing materials were completed anonymously. Test instructions were paraphrased by the test administrator, and procedural questions were solicited and answered. All participation, in keeping with school and Ethics Committee policies, was voluntary and no incentives were offered.

Instrumentation

The BDI is a 21-item scale that measures symptoms related to cognitive, behavioral, affective, and somatic components of depression. Although originally designed for use by trained interviewers, it is now most typically used as a self-report measure (Beck, Steer, & Garbin, 1988; Kearns, Cruickshaw, McGuigan, Riley, Shaw, & Snaith, 1982; Vredenburg, Krames, & Flett, 1985). For each item, respondents are presented with four

statements rated from 0 to 3 in terms of intensity, and asked to select the one which most accurately describes their own feelings; higher scores represent a more severe level of reported depression. Total scores range from 0 to 63 and are used to categorize four levels of depression: none to minimal (0-9), mild to moderate (10-18), moderate to severe (19-29), and severe (30-63) (Beck et al., 1988).

Baron and LaPlante (1984) have reported an internal consistency reliability of .80, and a test-retest reliability over an 8-week period of .74 for the BDI-FR for nonclinical adolescents. Substantially greater psychometric data have been reported for the original English version of the BDI with respect to the adolescent population (for a review, see Byrne & Baron, in press a; Byrne et al., 1991).

Analysis of the Data

Factorial validity of the BDI-FR was tested using analyses of covariance structures within the framework of the CFA model. For purposes of cross-validation, male and female data were each randomly split into two to form calibration and validation samples. Analyses were then conducted in three stages using the EQS program (Bentler, 1989). First, based on the calibration sample, CFA procedures were conducted separately for males and females to test the hypothesized hierarchical 3-factor structure

reported by Byrne and Baron (in press b). Given findings of inadequate fit, and sound statistical, empirical, and theoretical justification for doing so, the model was respecified to include additional parameters identified by the Lagrange Multiplier Test (LM-Test) as those that would contribute most to a significantly better-fitting model. Second, the final best-fitting model from Stage 1 was tested for its invariance across calibration and validation samples for each sex. Finally, the best-fitting model for males and females, as derived from Stage 1, was tested for its invariance across gender.

Assessment of model fit was based on multiple criteria that reflected statistical, theoretical, and practical considerations; these were: (a) the χ^2 likelihood ratio statistic, (b) the Comparative Fit Index (CFI; Bentler, 1990), (c) the Satorra-Bentler Scaled Statistic (S-Bχ^2; Satorra & Bentler, 1988), and (d) the substantive meaningfulness of the model (see MacCallum, 1986). The CFI is a revised version of the Bentler-Bonett (1980) normed fit index that adjusts for degrees of freedom. It ranges from zero to 1.00 and is derived from the comparison of a restricted model (i.e., one in which structure is imposed on the data) with a null model (one in which each observed variable represents a factor). The CFI provides a measure of complete covariation in the data; a value >.90 indicates a

psychometrically acceptable fit to the data. The S-Bχ^2 incorporates a scaling correction for the χ^2 statistic when distributional assumptions are violated. Its computation takes into account the model, the estimation method, and the sample kurtosis values (Hu, Bentler, & Kano, in press). The S-Bχ^2 has been shown to more closely approximate χ^2 than the usual test statistic, to have robust standard errors, and to perform as well or better than the usual asymptotically distribution-free (ADF) methods generally recommended for nonnormal multivariate data (Bentler, 1989; Hu et al., in press). Given the known kurtotic nature of particular BDI items for both English- and French speaking nonclinical adolescents (Byrne & Baron, in press a, in press b; Byrne et al., 1991), this statistic allows for a more cogent assessment of factorial validity than is possible with the uncorrected (i.e., biased) statistic. The corrected CFI value (CFI*) computed from the S-Bχ statistic for the null model is also reported. It is important to note, however, that the S-Bχ^2 is not yet available for multigroup analyses in the current version of the EQS program; these values are therefore reported only for the single-group analyses.

Results

Following from the earlier work of Byrne and Baron (in press), the CFA model in the present study hypothesized a priori

that: (a) responses to the BDI-FR could be explained by three first-order factors, and one second-order factor of General Depression, (b) each item would have a non-zero loading on the first-order factor it was designed to measure, and zero loadings on the other two first-order factors, (c) error terms associated with each item would be uncorrelated, and (d) covariation among the three first-order factors would be explained fully by their regression onto the second-order factor (see Figure 1).

Preliminary analyses identified three multivariate outliers for males, and one for females; all related to the calibration samples. Deletion of these cases resulted in calibration sample sizes of 273 and 300 for males and females, respectively. As expected, there was some evidence of multivariate positive kurtosis for both males and females as reflected by normalized Mardia coefficients of 58.22 and 45.52, respectively. While it is unlikely that the maximum likelihood estimates would be affected, nonnormality could lead to downwardly biased standard errors which would result in an inflated number of statistically significant parameters (Muthén & Kaplan, 1985). Thus, final assessment of statistical fit was based on the S-Bχ^2 which corrects for this violation, and on its related CFI* value.

Stage 1: Tests of the Hypothesized Model

Before testing the validity of the higher-order model purported to underlie the BDI-FR, it was important to test first for the psychometric adequacy of the lower-order structure. Given a well-fitting model, reparameterization of a second-order structure was in order and expected to exhibit an identical degree of fit. As indicated by the CFI* values reported in Table 1, goodness-of-fit for the initially hypothesized model of BDI-FR structure was exceptionally good for males; it was somewhat less well-fitting for females. We turn now to this problematic fit for the latter group.

Insert Table 1 about here

Examination of the multivariate LM χ^2 coefficients for females revealed substantial improvement in model fit to be gained from the additional specification Item 17 (fatigue) on the Negative Attitude factor, and four error covariances (positive covariances between Items 2 & 4, 8 & 5, and 7 & 14; negative covariance betweem Items 2 & 18). Thus, the initially hypothesized model was respecified to include these five parameters, and then reestimated.

To assess the extent to which a respecified model exhibits

an improvement in fit, we examine the difference in χ^2 ($\Delta\chi^2$) between the two models. This differential is itself χ^2-distributed, with degrees of freedom equal to the difference in degrees of freedom (Δdf) and can, thus, be tested statistically; a significant $\Delta\chi^2$ indicates a substantial improvement in model fit. As shown in Table 1, incorporation of these five parameters into the model resulted in a statistically better-fitting model for adolescent females (ΔS-B$\chi^2_{(5)}$=67.64); the difference in CFI* values was also substantial (.07).

Interestingly, although the cross-loading of Item 17 onto Factor 1 (Negative Attitude) was statistically significant ($p<.001$), as were the error covariances (all, $p<.01$), the initial loading of Item 17 onto Factor 3 (Somatic Elements) was not significant ($p=.08$). Thus, Model 1 was again reestimated with Item 17 loading onto Factor 1 instead of Factor 3, as originally hypothesized. Although the difference in S-Bχ^2 values between these two models was not statistically significant, overall fit for Model 4 was equivalent to that of Model 3 while being based on a more parsimonious simple structure that represented a psychometrically sounder parameterization of the instrument.

Once the lower-order factor structure was determined, the higher-order model was then specified. Estimation of this model revealed the disturbance term associated with Factor 1 to be a

boundary parameter which was subsequently fixed to .001; this accounted for the negligible difference in χ^2 values between Models 3 and 4. As evidenced by both the CFI and CFI* coefficients, this constraint led to no degradation in overall model fit. The final model of BDI-FR structure for males and females is presented schematically in Figure 2.

Insert Figure 2 about here

Stage 2: Tests for Equality Across Calibration/Validation Samples

To test the replication of these findings for males and females, the final model for each group was tested for its invariance across a second independent (i.e., validation) sample. As such, all first- and second-order factor loadings were constrained equal across calibration and validation samples, and then tested statistically in a simultaneous analysis of the data; for females, equality constraints were also specified for the four error covariances. Judgment of replicability was based on two criteria: (a) goodness-of-fit of the constrained model, and (b) probability level of the equality constraints as determined by the LM Test (equality constraints with $p<.05$ being untenable). In evaluating the fit of these models, readers are reminded that χ^2 values for the multigroup models are based on the true, rather

than on the S-Bχ^2 statistic. Given the known sample size dependency of the true χ^2 statistic, and the fact that it has not been scaled to correct for multivariate kurtosis, the χ^2 values are expected to be substantially larger than would be the case for the S-Bχ^2 statistic.

The previous caveat notwithstanding, results as shown in Table 2 revealed all constrained models to be well-fitting for both males and females, and all equality constraints to be tenable. Probability values associated with these constraints ranged from .09 to .97 (M=.63) for adolescent males, and from .61 to .99 (M=.85) for adolescent females.

Insert Table 2 about here

Stage 3: Tests for Equality Across Gender

Since the findings from Stage 2 of the analyses argued for the statistical equivalence of model structure across independent samples, calibration and validation samples were subsequently combined for males and for females. Tests for the invariance of factorial validity across gender were therefore based on full samples, less the multivariate outliers noted earlier (males, n = 548; females, n = 600). The focus of these analyses was to test for the equivalence of all first- and second-order factor

loadings across gender. Adhering to caveats bearing on partial measurement invariance, the gender-specific loading of Item #17 was left unconstrained across sex (see Byrne, Shavelson, & Muthen, 1989).

Remarkably, all equality constraints were found to be tenable across gender, and as indicated by the CFI values reported in Table 2, the invariant model was well-fitting. Probability values associated with these constraints ranged from .16 to .98 ($\underline{M}$=.76).

Discussion

Findings from this cross-validated study have contributed a wealth of valuable information to our growing knowledge of the BDI when used with nonclinical adolescents. In particular, it has provided critically needed validity data on the French version of the instrument which, to date, has been sadly lacking. Indeed, in light of the rigorous testing procedures used in this study, the factorial structure of the BDI-FR, as originally described by Byrne and Baron (in press b), has proven to be remarkably psychometrically sound in its separate measurement of depressic for French-speaking adolescent males and females. Except for one item (#17; fatigue), all lower- and higher-order factor loadings were found to be equivalent across gender. Within gender, all factor loadings replicated across calibration and validation

samples; for females, replication of the additionally specified error covariances was also tenable.

While the differential loading of Item 17 onto the Negative Attitude factor, rather than the Somatic Elements factor for females is intriguing from a gender difference perspective, it speaks to inquiry that also crosses cultural boundaries. For example, in their initial validation of the BDI for English-speaking adolescents, Byrne and Baron (in press a) found this item to load significantly and substantially onto the Performance Difficulty factor. Moreover, this specification was found to be invariant both across independent validation samples and across gender (Byrne et al., 1991). In testing the validity of BDI structure (based on English adolescent data) as it related to French-speaking adolescents, Byrne and Baron (in press b) found Item 17 to load significantly and substantially on Somatic Elements, rather than on the Performance Difficulty factor. Our present findings now suggest that for French-speaking adolescents, the loading of this item differs for males and females.

Careful scrutiny of the fatigue item suggests that the wording may be ambiguous for both French and English versions of the instrument. More specifically, endorsement at lower levels seems more reflective of somatic concerns (e.g., don't get more

tired than usual") whereas endorsement at a higher level may be more indicative of motivational concerns or lack of mental energy (e.g., I am too tired to do anything"). Indeed, fatigue could conceivably load, and does, on any of the first-order factors. Thus, fatigue may be a weak item, the interpretation of which should be carefully monitored if subscale scores are to be used for clinical or research purposes.

Our finding of error covariances for females, and not for males also poses an interesting question that appears linked to cultural, as well as gender issues. While these parameters were not apparent for English-speaking adolescents in general (Byrne & Baron, in press a), one such parameter was identified for females (Items 20 & 21) when the issue of gender was taken into account (Byrne et al., 1991). Error covariances were also a significant presence in the validation of the BDI for French-speaking adolescents in general (Byrne & Baron, in press b).

Error covariances most often represent nonrandom measurement error due to method effects associated with the response format of measuring instruments, and are not unexpected in the CFA of a single assessment scale (see e.g., Byrne, 1988, in press; Tanaka & Huba, 1984). The association between Items 2 (pessimism) and 4 (dissatisfaction) is possibly explained within the context of the hopelessness theory of depression (Metalsley & Alloy, 1989). As

such, pessimism may contribute to symptoms of hopelessness depression (e.g., motivational deficits). Item 4 seems conceptually linked to motivational concerns. In view of research suggesting that males exhibit higher levels of internal locus of control than females (Siddique & D'Arcy, 1984), this conceptualization may be more representative of adolescent females than their male peers. The error covariance between Items 7 (self-dislike) and 14 (self/body image distortion) may reflect the observation that boys are generally more satisfied with their bodies than are girls (Rierdan, Koff, & Stubbs, 1988); poor body image may, in turn, contribute to feelings of self-dislike.

That error covariances contributed significantly to model misfit for females, and not for males, in validating both the English and French versions of the BDI is intriguing. However, their differential specification across Anglophone and Francophone females is puzzling, as is failure of the three remaining error covariances (Items 2 & 3, Byrne & Baron, in press b; Items 8 & 5, Items 2 & 18, present study) to replicate across studies. One possible explanation of this finding may lie with the differential algorithmic approach to assessment of model fit by the LISREL and EQS programs used in the former and present studies, respectively. More specifically, whereas LISREL uses a univariate approach in pinpointing model misfit, EQS uses a

multivariate approach.

Of final import, from a gender perspective, is the higher loading of Item 10 (crying) for females than for males onto the Negative Attitude factor. Interestingly, the same finding was demonstrated for Anglophone females (Byrne et al., 1991), and is consistent with research that has shown crying to be more characteristic of females than it is of males (Campbell, Byrne, & Baron, 1992; Gruchy & Stoppard, 1992).

Except for one differential factor loading and two error covariances, results provide strong support for the psychometric equality of the BDI-FR across adolescent males and females. As with all scientific inquiry, however, substantially more replication work is needed in order to confirm or disconfirm our findings. Given (a) the increasing incidence of adolescent depression, and (b) the sparseness of statistically stringent validity research on both the English and French versions of the BDI as it bears on this population, we hope that our work will encourage others to follow suit thereby adding to our growing knowledge of BDI factorial structure relative to nonclinical contexts.

References

Abramson, L.Y., Metalsky, G.J., & Alloy, L.B. (1989). Hopelessness depression: A theory-based subtype of depression. Psychological Review, 96, 358-372.

Baron, P., & Laplante, L. (1984). L'Inventaire de Dépression de Beck : Son utilisation auprès d'un échantillon d'adolescents francophones. Revue de Modification du Comportement, 14, 161-166.

Beck, A.T., Steer, R.A., & Garbin, M.G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical Psychology Review, 8, 77-100.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4, 561-571.

Bentler, P.M. (1989). EQS: Structural Equations program manual. Los Angeles, CA: BMDP Statistical Software.

Bentler, P.M. (1990). Comparative fit indices in structural models. Psychological Bulletin, 107, 238-246.

Bentler, P.M. & Bonett, D.G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. Psychological Bulletin, 88, 588-606.

Bourque, P. & Beaudette, D. (1982). Étude psychométrique du questionnaire de dépression de Beck auprès d'un échantillon d'étudiants universitaires francophones. Revue canadienne des sciences du comportement, 14, 211-21.

Byrne, B.M. (1988). Measuring adolescent self-concept: Factorial validity and equivalency of the SDQ III across gender. Multivariate Behavioral Research, 23, 361-375.

Byrne, B.M. (in press). The Maslach Burnout Inventory: Testing for factorial validity and invariance across elementary, intermediate, and secondary teachers. Journal of Occupational and Organizational Psychology.

Byrne, B.M. & Baron, P. (in press a). The Beck Depression Inventory: Testing and cross-validating an hierarchical factor structure for nonclinical adolescents. Measurement and Development in Counseling and Development.

Byrne, B.M. & Baron, P. (in press b). Measuring adolescent depression: Tests of equivalent factorial structure for English and French versions of the Beck Depression Inventory. Applied Psychology: An International Review.

Byrne, B.M., Campbell, T.L., & Baron, P. (1991, August). The Beck Depression Inventory: Testing for invariant measurement and structure across gender for nonclinical adolescents. Paper presented at the American Psychological Association Annual Meeting, San Francisco. ERIC Document Reproduction Services No. TM017089.

Byrne, B.M., Shavelson, R.J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. Psychological Bulletin, 105, 456-466.

Campbell, T.L., Byrne, B.M., & Baron, P. (1992). Gender differences in the expression of depressive symptoms in early adolescents. Journal of Early Adolescence, 12, 326-338.

Gruchy, C.A. & Stoppard, J.M. (1992, June). Gender, context, and crying as an expression of sadness. Poster presented at the Canadian Psychological Annual Meeting, Quebec City.

Hu, L-, Bentler, P.M., & Kano, Y. (in press). Can test statistics in covariance structure analysis be trusted? Psychological Bulletin.

Kearns, N.P., Cruickshank, C.A., McGuigan, K.J., Riley, S.A., Shaw, S.P., & Snaith, R.P. (1982). A comparison of depression rating scales. British Journal of Psychiatry, 141, 45-49.

MacCallum, R. (1986). Specification searches in covariance structure modeling. Psychological Bulletin, 100, 107-120.

Muthen, B. & Kaplan, D. (1985). A comparison of methodologies for the factor analysis of non-normal Likert variables. British Journal of Mathematical and Statistical Psychology, 38, 171-189.

Rierdan, J., Koff, E., & Stubbs, M.L. (1987). Depressive symptomatology and body image in adolescent girls. Journal of Early Adolescence, 8, 109-117.

Satorra, A. & Bentler, P.M. (1988). Scaling correcctions for chi-square in covariance structure analysis. Proceedings of the American Statistical Association.

Siddique, E.M. & D'Arcy, C. (1984). Adolescence, stress, and psychological well-being. Journal of Youth and Adolescence, 13, 459-473.

Tanaka, J.S. & Huba, G.J. (1984). Confirmatory hierarchical factor analyses of psychological distress measures. Journal of Personality and Social Psychology, 46, 621-635.

Vredenburg, K., Krames, L., & Flett, G.L. (1985). Reexamining the Beck Depression Inventory: The long and short of it. Psychological Reports, 56, 767-778.

Table 1

Goodness-of-fit Statistics for Male and Female Calibration Samples

Model	χ^2	df	CFI	S-Bχ^2	CFI*[a]
		Males			
1 Initial (1st-order)	285.66	186	.85	212.86	.94
2 Model 1 (2nd-order)	285.66	186	.85	212.86	.94
		Females			
1 Initial (1st-order)	393.51	186	.85	296.99	.88
2 Model 1 with: •Item 17 cross-loaded on F1 •4 correlated errors	297.25	181	.92	229.35	.95
3 Model 1 with: •Item 17 loaded on F1 instead of on F3 •4 correlated errors	301.18	182	.92	232.29	.95
4 Model 3 (2nd-order)	307.52	183[b]	.92	236.75	.95

[a] corrected Comparative Fit Index based on Satorra-Bentler χ^2 fit for the null model

[b] disturbance term for F1 fixed to .001

F1 = Factor 1 (Negative Attitude); F3 = Factor 3 (Somatic Elements)

Table 2

Tests for Measurement Invariance Within and Across Gender

Model	χ^2	df	CFI	LM[a] Equality Constraint Statistics
1 All factor loadings[b] invariant across calibration (n=272) and validation (n=275) male samples	608.41	393	.86	All equality constraints tenable (p>.09)
2 All factor loadings[b] and correlated errors invariant across calibration (n=300) and validation (n=300) females samples	623.81	389	.92	All equality constraints tenable (p>.61)
3 All factor loadings[b], except the loading of Item 17 on F1, invariant across males (n=548) and females (n=600)	659.01	388	.93	All equality constraints tenable (p>.16)

[a]Lagrange Multiplier

[b]First- and second-order loadings

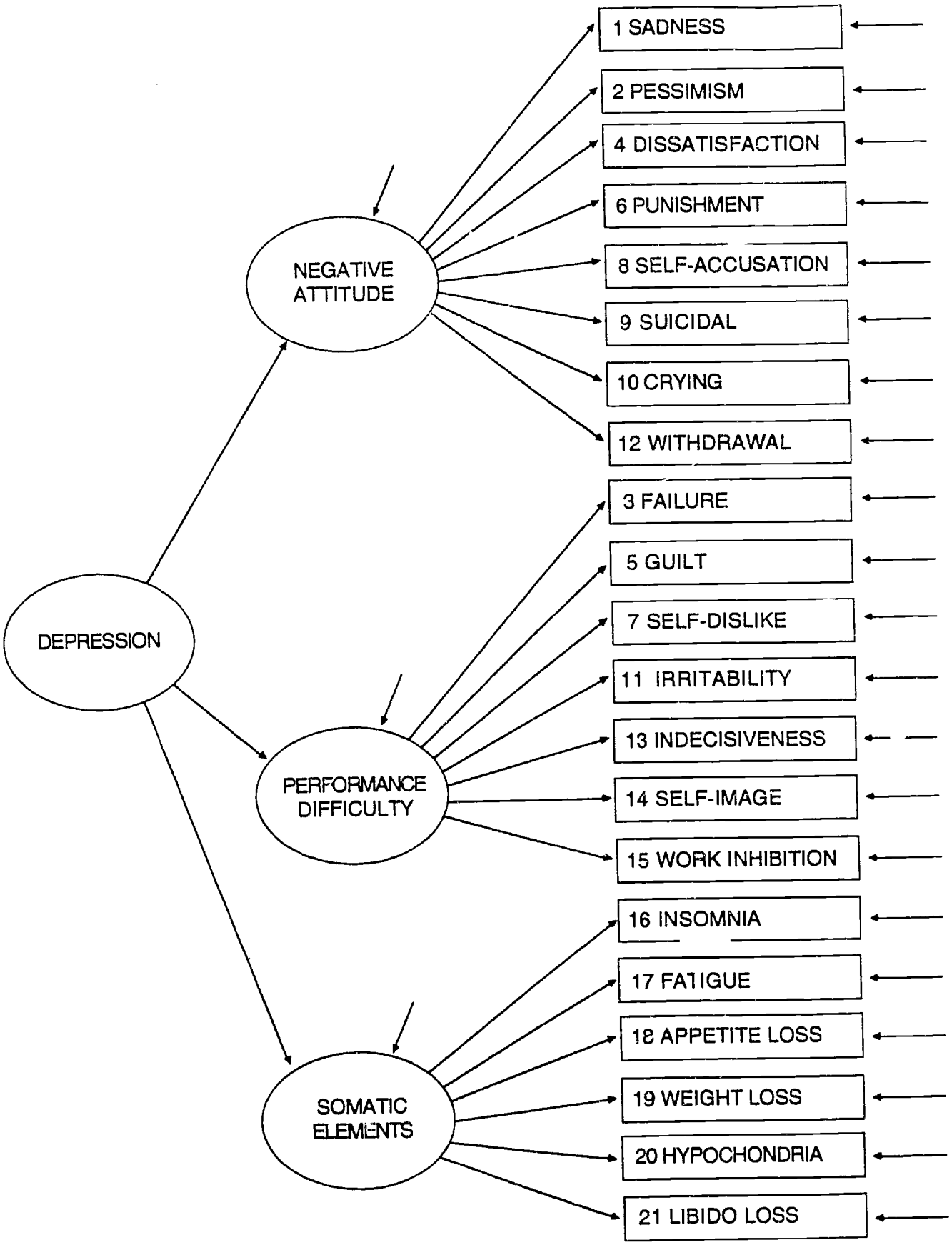
F1 = Factor 1 (Negative Attitude)

Figure Captions

Figure 1. Hypothesized Model of BDI Structure for French-speaking Nonclinical Adolescents

Figure 2. Standardized estimates for final models of BDI structure (French version) for nonclinical adolescent males and females. Parenthesized values represent S-Bχ^2-corrected z-scores. Values in boxes represent item numbers.

* denotes parameter fixed to 1.0 for purposes of statistical identification


DEPRESSION
NEGATIVE ATTITUDE
PERFORMANCE DIFFICULTY
SOMATIC ELEMENTS
1 SADNESS
2 PESSIMISM
4 DISSATISFACTION
6 PUNISHMENT
8 SELF-ACCUSATION
9 SUICIDAL
10 CRYING
12 WITHDRAWAL
3 FAILURE
5 GUILT
7 SELF-DISLIKE
11 IRRITABILITY
13 INDECISIVENESS
14 SELF-IMAGE
15 WORK INHIBITION
16 INSOMNIA
17 FATIGUE
18 APPETITE LOSS
19 WEIGHT LOSS
20 HYPOCHONDRIA
21 LIBIDO LOSS